

RAW SEQUENCE LISTING

**The Biotechnology Systems Branch of the Scientific and Technical
Information Center (STIC) no errors detected.**

Application Serial Number: 10/511, 718
Source: PCT
Date Processed by STIC: 10/26/2005

ENTERED



PCT

RAW SEQUENCE LISTING

DATE: 10/26/2005

PATENT APPLICATION: US/10/511,718

TIME: 13:46:33

Input Set : A:\pto.da.txt

Output Set: N:\CRF4\10262005\J511718.raw

3 <110> APPLICANT: Innovations Foundation

5 <120> TITLE OF INVENTION: Recombinant BCG Strains Expressing Alanine Dehydrogenase,

Serine

6 dehydratase and/or Glutamine Synthetase as TB Vaccines

W--> 8 <130> FILE REFERENCE:

C--> 10 <140> CURRENT APPLICATION NUMBER: US/10/511,718

C--> 10 <141> CURRENT FILING DATE: 2004-10-18

10 <150> PRIOR APPLICATION NUMBER: US 60/372,450

11 <151> PRIOR FILING DATE: 2002-04-16

13 <160> NUMBER OF SEQ ID NOS: 14

15 <170> SOFTWARE: PatentIn version 3.0

17 <210> SEQ ID NO: 1

18 <211> LENGTH: 1116

19 <212> TYPE: DNA

20 <213> ORGANISM: Mycobacterium tuberculosis

22 <220> FEATURE:

23 <221> NAME/KEY: CDS

24 <222> LOCATION: (1)..(1116)

25 <223> OTHER INFORMATION: Sequence is identical to GenBank entries GI:3089350 and

MTU92472

28 <400> SEQUENCE: 1

29 atg cgc gtc ggt att ccg acc gag acc aaa aac aac gaa ttc cgg gtg 48

30 Met Arg Val Gly Ile Pro Thr Glu Thr Lys Asn Asn Glu Phe Arg Val

31 1 5 10 15

33 gcc atc acc ccg gcc ggc gtc gcg gaa cta acc cgt cgt ggc cat gag 96

34 Ala Ile Thr Pro Ala Gly Val Ala Glu Leu Thr Arg Arg Gly His Glu

35 20 25 30

37 gtg ctc atc cag gca ggt gcc gga gag ggc tcg gct atc acc gac gcg 144

38 Val Leu Ile Gln Ala Gly Ala Gly Glu Gly Ser Ala Ile Thr Asp Ala

39 35 40 45

41 gat ttc aag gcg gca ggc gcg caa ctg gtc ggc acc gcc gac cag gtg 192

42 Asp Phe Lys Ala Ala Gly Ala Gln Leu Val Gly Thr Ala Asp Gln Val

43 50 55 60

45 tgg gcc gac gct gat tta ttg ctc aag gtc aaa gaa ccg ata gcg gcg 240

46 Trp Ala Asp Ala Asp Leu Leu Leu Lys Val Lys Glu Pro Ile Ala Ala

47 65 70 75 80

49 gaa tac ggc cgc ctg cga cac ggg cag atc ttg ttc acg ttc ttg cat 288

50 Glu Tyr Gly Arg Leu Arg His Gly Gln Ile Leu Phe Thr Phe Leu His

51 85 90 95

53 ttg gcc gcg tca cgt gct tgc acc gat gcg ttg ttg gat tcc ggc acc 336

54 Leu Ala Ala Ser Arg Ala Cys Thr Asp Ala Leu Leu Asp Ser Gly Thr

55 100 105 110

57 acg tca att gcc tac gag acc gtc cag acc gcc gac ggc gca cta ccc 384

58 Thr Ser Ile Ala Tyr Glu Thr Val Gln Thr Ala Asp Gly Ala Leu Pro

59

115

120

125

RAW SEQUENCE LISTING

DATE: 10/26/2005

PATENT APPLICATION: US/10/511,718

TIME: 13:46:33

Input Set : A:\pto.da.txt

Output Set: N:\CRF4\10262005\J511718.raw

```

61 ctg ctt gcc ccg atg agc gaa gtc gcc ggt cga ctc gcc gcc cag gtt      432
62 Leu Leu Ala Pro Met Ser Glu Val Ala Gly Arg Leu Ala Ala Gln Val
63      130                      135                      140
65 ggc gct tac cac ctg atg cga acc caa ggg ggc cgc ggt gtg ctg atg      480
66 Gly Ala Tyr His Leu Met Arg Thr Gln Gly Gly Arg Gly Val Leu Met
67 145                      150                      155                      160
69 ggc ggg gtg ccc ggc gtc gaa ccg gcc gac gtc gtg gtg atc ggc gcc      528
70 Gly Gly Val Pro Gly Val Glu Pro Ala Asp Val Val Val Ile Gly Ala
71      165                      170                      175
73 ggc acc gcc ggc tac aac gca gcc cgc atc gcc aac ggc atg ggc gcg      576
74 Gly Thr Ala Gly Tyr Asn Ala Ala Arg Ile Ala Asn Gly Met Gly Ala
75      180                      185                      190
77 acc gtt acg gtt cta gac atc aac atc gac aaa ctt cgg caa ctc gac      624
78 Thr Val Thr Val Leu Asp Ile Asn Ile Asp Lys Leu Arg Gln Leu Asp
79      195                      200                      205
81 gcc gag ttc tgc ggc cgg atc cac act cgc tac tca tcg gcc tac gag      672
82 Ala Glu Phe Cys Gly Arg Ile His Thr Arg Tyr Ser Ser Ala Tyr Glu
83      210                      215                      220
85 ctc gag ggt gcc gtc aaa cgt gcc gac ctg gtg att ggg gcc gtc ctg      720
86 Leu Glu Gly Ala Val Lys Arg Ala Asp Leu Val Ile Gly Ala Val Leu
87 225                      230                      235                      240
89 gtg cca ggc gcc aag gca ccc aaa tta gtc tcg aat tca ctt gtc gcg      768
90 Val Pro Gly Ala Lys Ala Pro Lys Leu Val Ser Asn Ser Leu Val Ala
91      245                      250                      255
93 cat atg aaa cca ggt gcg gta ctg gtg gat ata gcc atc gac cag ggc      816
94 His Met Lys Pro Gly Ala Val Leu Val Asp Ile Ala Ile Asp Gln Gly
95      260                      265                      270
97 ggc tgt ttc gaa ggc tca cga ccg acc acc tac gac cac ccg acg ttc      864
98 Gly Cys Phe Glu Gly Ser Arg Pro Thr Thr Tyr Asp His Pro Thr Phe
99      275                      280                      285
101 gcc gtg cac gac acg ctg ttt tac tgc gtg gcg aac atg ccc gcc tcg      912
102 Ala Val His Asp Thr Leu Phe Tyr Cys Val Ala Asn Met Pro Ala Ser
103      290                      295                      300
105 gtg ccg aag acg tcg acc tac gcg ctg acc aac gcg acg atg ccg tat      960
106 Val Pro Lys Thr Ser Thr Tyr Ala Leu Thr Asn Ala Thr Met Pro Tyr
107 305                      310                      315                      320
109 gtg ctc gag ctt gcc gac cat ggc tgg cgg gcg gcg tgc cgg tcg aat      1008
110 Val Leu Glu Leu Ala Asp His Gly Trp Arg Ala Ala Cys Arg Ser Asn
111      325                      330                      335
113 ccg gca cta gcc aaa ggt ctt tcg acg cac gaa ggg gcg tta ctg tcc      1056
114 Pro Ala Leu Ala Lys Gly Leu Ser Thr His Glu Gly Ala Leu Leu Ser
115      340                      345                      350
117 gaa cgg gtg gcc acc gac ctg ggg gtg ccg ttc acc gag ccc gcc agc      1104
118 Glu Arg Val Ala Thr Asp Leu Gly Val Pro Phe Thr Glu Pro Ala Ser
119      355                      360                      365
121 gtg ctg gcc tga      1116
122 Val Leu Ala
123      370
126 <210> SEQ ID NO: 2

```

RAW SEQUENCE LISTING

DATE: 10/26/2005

PATENT APPLICATION: US/10/511,718

TIME: 13:46:33

Input Set : A:\pto.da.txt

Output Set: N:\CRF4\10262005\J511718.raw

127 <211> LENGTH: 371
 128 <212> TYPE: PRT
 129 <213> ORGANISM: Mycobacterium tuberculosis
 131 <220> FEATURE:
 W--> 132 <221> NAME/KEY:
 133 <222> LOCATION:
 134 <223> OTHER INFORMATION: Sequence is identical to SwissProt entry SP:DHA_MYCTU
 136 <400> SEQUENCE: 2
 137 Met Arg Val Gly Ile Pro Thr Glu Thr Lys Asn Asn Glu Phe Arg Val
 138 1 5 10 15
 141 Ala Ile Thr Pro Ala Gly Val Ala Glu Leu Thr Arg Arg Gly His Glu
 142 20 25 30
 145 Val Leu Ile Gln Ala Gly Ala Gly Glu Gly Ser Ala Ile Thr Asp Ala
 146 35 40 45
 149 Asp Phe Lys Ala Ala Gly Ala Gln Leu Val Gly Thr Ala Asp Gln Val
 150 50 55 60
 153 Trp Ala Asp Ala Asp Leu Leu Leu Lys Val Lys Glu Pro Ile Ala Ala
 154 65 70 75 80
 157 Glu Tyr Gly Arg Leu Arg His Gly Gln Ile Leu Phe Thr Phe Leu His
 158 85 90 95
 161 Leu Ala Ala Ser Arg Ala Cys Thr Asp Ala Leu Leu Asp Ser Gly Thr
 162 100 105 110
 165 Thr Ser Ile Ala Tyr Glu Thr Val Gln Thr Ala Asp Gly Ala Leu Pro
 166 115 120 125
 169 Leu Leu Ala Pro Met Ser Glu Val Ala Gly Arg Leu Ala Ala Gln Val
 170 130 135 140
 173 Gly Ala Tyr His Leu Met Arg Thr Gln Gly Gly Arg Gly Val Leu Met
 174 145 150 155 160
 177 Gly Gly Val Pro Gly Val Glu Pro Ala Asp Val Val Val Ile Gly Ala
 178 165 170 175
 181 Gly Thr Ala Gly Tyr Asn Ala Ala Arg Ile Ala Asn Gly Met Gly Ala
 182 180 185 190
 185 Thr Val Thr Val Leu Asp Ile Asn Ile Asp Lys Leu Arg Gln Leu Asp
 186 195 200 205
 189 Ala Glu Phe Cys Gly Arg Ile His Thr Arg Tyr Ser Ser Ala Tyr Glu
 190 210 215 220
 193 Leu Glu Gly Ala Val Lys Arg Ala Asp Leu Val Ile Gly Ala Val Leu
 194 225 230 235 240
 197 Val Pro Gly Ala Lys Ala Pro Lys Leu Val Ser Asn Ser Leu Val Ala
 198 245 250 255
 201 His Met Lys Pro Gly Ala Val Leu Val Asp Ile Ala Ile Asp Gln Gly
 202 260 265 270
 205 Gly Cys Phe Glu Gly Ser Arg Pro Thr Thr Tyr Asp His Pro Thr Phe
 206 275 280 285
 209 Ala Val His Asp Thr Leu Phe Tyr Cys Val Ala Asn Met Pro Ala Ser
 210 290 295 300
 213 Val Pro Lys Thr Ser Thr Tyr Ala Leu Thr Asn Ala Thr Met Pro Tyr
 214 305 310 315 320
 217 Val Leu Glu Leu Ala Asp His Gly Trp Arg Ala Ala Cys Arg Ser Asn

RAW SEQUENCE LISTING

DATE: 10/26/2005

PATENT APPLICATION: US/10/511,718

TIME: 13:46:33

Input Set : A:\pto.da.txt

Output Set: N:\CRF4\10262005\J511718.raw

```

218          325          330          335
221 Pro Ala Leu Ala Lys Gly Leu Ser Thr His Glu Gly Ala Leu Leu Ser
222          340          345          350
225 Glu Arg Val Ala Thr Asp Leu Gly Val Pro Phe Thr Glu Pro Ala Ser
226          355          360          365
229 Val Leu Ala
230          370
233 <210> SEQ ID NO: 3
234 <211> LENGTH: 399
235 <212> TYPE: DNA
236 <213> ORGANISM: Mycobacterium bovis
238 <220> FEATURE:
239 <221> NAME/KEY: CDS
240 <222> LOCATION: (1)..(399)
242 <400> SEQUENCE: 3
243 atg cgc gtc ggt att ccg acc gag acc aaa aac aac gaa ttc cgg gtg      48
244 Met Arg Val Gly Ile Pro Thr Glu Thr Lys Asn Asn Glu Phe Arg Val
245 1          5          10          15
247 gcc atc acc ccg gcc ggc gtc gcg gaa cta acc cgt cgt ggc cat gag      96
248 Ala Ile Thr Pro Ala Gly Val Ala Glu Leu Thr Arg Arg Gly His Glu
249          20          25          30
251 gtg ctc atc cag gca ggt gcc gga gag ggc tcg gct atc acc gac gcg      144
252 Val Leu Ile Gln Ala Gly Ala Gly Glu Gly Ser Ala Ile Thr Asp Ala
253          35          40          45
255 gat ttc aag gcg gca ggc gcg caa ctg gtc ggc acc gcc gac cag gtg      192
256 Asp Phe Lys Ala Ala Gly Ala Gln Leu Val Gly Thr Ala Asp Gln Val
257          50          55          60
259 tgg gcc gac gct gat tta ttg ctc aag gtc aaa gaa ccg ata gcg gcg      240
260 Trp Ala Asp Ala Asp Leu Leu Leu Lys Val Lys Glu Pro Ile Ala Ala
261 65          70          75          80
263 gaa tac ggc cgc ctg cga cac ggg cga tct tgt tca cgt tct tgc att      288
264 Glu Tyr Gly Arg Leu Arg His Gly Arg Ser Cys Ser Arg Ser Cys Ile
265          85          90          95
267 tgg ccg cgt cac gtg ctt gca ccg atg cgt tgt tgg att ccg gca cca      336
268 Trp Pro Arg His Val Leu Ala Pro Met Arg Cys Trp Ile Pro Ala Pro
269          100          105          110
271 cgt caa ttg cct acg aga ccg tcc aga ccg ccg acg gcg cac tac ccc      384
272 Arg Gln Leu Pro Thr Arg Pro Ser Arg Pro Pro Thr Ala His Tyr Pro
273          115          120          125
275 tgc ttg ccc cga tga      399
276 Cys Leu Pro Arg
277          130
280 <210> SEQ ID NO: 4
281 <211> LENGTH: 132
282 <212> TYPE: PRT
283 <213> ORGANISM: Mycobacterium bovis
285 <400> SEQUENCE: 4
286 Met Arg Val Gly Ile Pro Thr Glu Thr Lys Asn Asn Glu Phe Arg Val
287 1          5          10          15

```

RAW SEQUENCE LISTING

DATE: 10/26/2005

PATENT APPLICATION: US/10/511,718

TIME: 13:46:33

Input Set : A:\pto.da.txt

Output Set: N:\CRF4\10262005\J511718.raw

```

290 Ala Ile Thr Pro Ala Gly Val Ala Glu Leu Thr Arg Arg Gly His Glu
291          20          25          30
294 Val Leu Ile Gln Ala Gly Ala Gly Glu Gly Ser Ala Ile Thr Asp Ala
295          35          40          45
298 Asp Phe Lys Ala Ala Gly Ala Gln Leu Val Gly Thr Ala Asp Gln Val
299          50          55          60
302 Trp Ala Asp Ala Asp Leu Leu Leu Lys Val Lys Glu Pro Ile Ala Ala
303 65          70          75          80
306 Glu Tyr Gly Arg Leu Arg His Gly Arg Ser Cys Ser Arg Ser Cys Ile
307          85          90          95
310 Trp Pro Arg His Val Leu Ala Pro Met Arg Cys Trp Ile Pro Ala Pro
311          100         105         110
314 Arg Gln Leu Pro Thr Arg Pro Ser Arg Pro Pro Thr Ala His Tyr Pro
315          115         120         125
318 Cys Leu Pro Arg
319          130

```

322 <210> SEQ ID NO: 5

323 <211> LENGTH: 1386

324 <212> TYPE: DNA

325 <213> ORGANISM: Mycobacterium tuberculosis

327 <220> FEATURE:

328 <221> NAME/KEY: CDS

329 <222> LOCATION: (1)..(1386)

330 <223> OTHER INFORMATION: Sequence is identical to the complement of nucleotides

13172-14551

331 of GenBank entry GB:MTV030 [AL021428]

332 Sequence is identical to the complement of nucleotides 13195-14580

333 of GenBank entry GB:AE006919

336 <400> SEQUENCE: 5

```

337 atg acc atc agc gtc ttc gac ctg ttc acc atc ggc atc ggg ccg tcc      48
338 Met Thr Ile Ser Val Phe Asp Leu Phe Thr Ile Gly Ile Gly Pro Ser
339 1          5          10          15
341 agt tcc cac acc gtg gga ccg atg cgc gcg gca aac cag ttc gta gtt      96
342 Ser Ser His Thr Val Gly Pro Met Arg Ala Ala Asn Gln Phe Val Val
343          20          25          30
345 gcg ctg cgc cgc cgg ggc cac ctg gat gac ctc gag gcg atg cga gtg      144
346 Ala Leu Arg Arg Arg Gly His Leu Asp Asp Leu Glu Ala Met Arg Val
347          35          40          45
349 gat ctg ttc ggc tcg ctc gcg gcc acc gga gcc ggt cat ggc acc atg      192
350 Asp Leu Phe Gly Ser Leu Ala Ala Thr Gly Ala Gly His Gly Thr Met
351          50          55          60
353 tcg gcg ata ttg ctg ggg ctg gaa ggc tgc cag cca gaa acg att acc      240
354 Ser Ala Ile Leu Leu Gly Leu Glu Gly Cys Gln Pro Glu Thr Ile Thr
355 65          70          75          80
357 acc gaa cac aag gaa cgc cgg ctc gcc gag atc gca gcg tcc ggc gtg      288
358 Thr Glu His Lys Glu Arg Arg Leu Ala Glu Ile Ala Ala Ser Gly Val
359          85          90          95
361 acg cga atc ggc ggt gtc att ccg gtc ccg ctg acc gag cgt gat atc      336
362 Thr Arg Ile Gly Gly Val Ile Pro Val Pro Leu Thr Glu Arg Asp Ile
363          100         105         110

```

RAW SEQUENCE LISTING ERROR SUMMARY

DATE: 10/26/2005

PATENT APPLICATION: US/10/511,718

TIME: 13:46:34

Input Set : A:\pto.da.txt

Output Set: N:\CRF4\10262005\J511718.raw

Invalid Line Length:

The rules require that a line not exceed 72 characters in length. This includes spaces.

Seq#:7; Line(s) 592,594

VERIFICATION SUMMARY

DATE: 10/26/2005

PATENT APPLICATION: US/10/511,718

TIME: 13:46:34

Input Set : A:\pto.da.txt

Output Set: N:\CRF4\10262005\J511718.raw

L:8 M:201 W: Mandatory field data missing, <130> FILE REFERENCE
L:10 M:270 C: Current Application Number differs, Replaced Current Application No
L:10 M:271 C: Current Filing Date differs, Replaced Current Filing Date
L:132 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:2
L:460 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:6
L:725 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:8
L:986 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:10
L:1243 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:12
L:1503 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:14